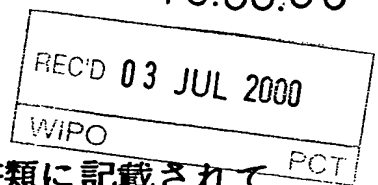


日本国特許庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

PCT/JP 00/03113

16.05.00



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日  
Date of Application:

1999年 5月21日

出願番号  
Application Number:

平成11年特許願第141838号

出願人  
Applicant(s):

松下電器産業株式会社

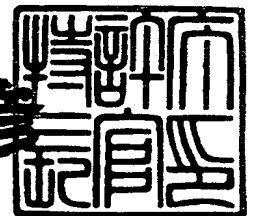
09/743578

PRIORITY  
DOCUMENT  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

2000年 6月16日

特許庁長官  
Commissioner,  
Patent Office

近藤 隆彦



出証番号 出証特2000-3045091

【書類名】 特許願

【整理番号】 2110011081

【提出日】 平成11年 5月21日

【あて先】 特許庁長官殿

【国際特許分類】 H04S 3/00

H04N 5/60

【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式会社内

【氏名】 小田 幹夫

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100097445

【弁理士】

【氏名又は名称】 岩橋 文雄

【選任した代理人】

【識別番号】 100103355

【弁理士】

【氏名又は名称】 坂口 智康

【選任した代理人】

【識別番号】 100109667

【弁理士】

【氏名又は名称】 内藤 浩樹

【手数料の表示】

【予納台帳番号】 011305

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9809938

【プルーフの要否】 不要

【書類名】 明細書

【発明の名称】 音声認識装置

【特許請求の範囲】

【請求項 1】 入力された音声信号の周波数成分を分析し、入力音声を音声認識データの標準音程に音程変換することで、音声認識率の向上を得る音声認識装置。

【請求項 2】 入力された音声信号をアナログ信号からデジタル信号に変換する A/D コンバータと、前記 A/D コンバータでデジタル信号に変換された入力音声信号を音程変換する音程変換処理部と、前記音程変換処理部で音程変換された音声デジタル信号を基に、音声認識分析する音声認識処理部を直列に具備し、前記音声認識処理部と並列に、音声周波数成分パターンを参照する標準音声データ部を具備すると共に、前記 A/D コンバータの出力を入力とし、入力音声信号の周波数成分を分析する周波数成分分析部と、前記周波数成分分析部の分析結果を入力とし、入力された音声信号の音程をどの程度上下に音程変換するかを判断する音程変換制御部を直列に具備し、前記音程変換制御部で決定された音程変換率で前記音程変換処理部により音程変換を行なう構成としたことを特徴とする音声認識装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、不特定話者の音声認識装置において、低音の男性の声や、高音の女性、子供の声にも幅広く音声認識処理が可能な音声認識装置に関するものである。

【0002】

【従来の技術】

昨今の音声認識技術は、デジタル信号処理技術の向上、処理 L S I の高性能化、低価格化などにより、民生機器に数多く導入が図られており、機器の操作性向上に役立っている。しかし、民生機器では、不特定話者を音声入力として認識しなければならず、話者は色々な話し方で入力する。例えば、早口でしゃべる人、

低音の男性、高音の女性、子供など、様々な特徴があり、これらの全てに対応し、高い音声認識率を実現するには課題も多い。これらの課題を解決する音声認識装置として、特開平9-325798に示す構成が提案されている。

## 【0003】

以下図4を参照しながら音声認識装置の一例について説明する。図4において、符号11は発声された音声信号を入力し、アナログ信号からデジタル信号に変換A/D変換する音声入力部、12は入力音声の発声速度を算出する発声速度算出部、13は発声速度算出部で算出された発声速度を基準速度と比較し、速度変換率の決定を行なう発声速度変換率決定部、14はその速度変換率を基に発声速度を変換する発声速度変換部、15は速度変換された入力音声信号を音声認識する音声認識部である。

## 【0004】

以上のように構成された音声認識装置について、その動作を説明する。発声入力された音声信号は音声入力部11でマイク、増幅器を経由して、A/Dコンバータにより、アナログ信号からデジタル信号に変換され、発声速度算出部12で入力音声の単語一音を切り出し、その一音の切り出し時間より、一音の発声速度を算出する。今、その一音の切り出し所要時間を $T_s$ とする。

## 【0005】

通常、一音を発声する基準時間を $T_h$ とし、発声速度変換率決定部13で発声速度算出部で算出された発声速度 $1/T_s$ を、基準速度 $1/T_h$ と比較し、速度変換率の決定を行なう。速度変換率は $T_s/T_h$ で求められ、 $T_s$ が $T_h$ より短ければ速度変換率は1以下、つまり遅くする。逆に $T_s$ が $T_h$ より長ければ速度変換率は1以上となり、早くすることを意味する。

## 【0006】

この様にして求められた速度変換率により、発声速度変換部14でその速度変換率を基に発声速度を変換し、一定の発声速度に速度変換し、速度変換された入力音声信号を、音声認識部15で音声認識処理され認識結果が出力される。速度変換は最近のデジタル技術を使用すれば、簡単に可能であり、速度を遅くする場合は、その一音の相関のある母音波形を複数追加して延ばし、早くする場合はそ

の一音の母音波形を複数間引くことにより実現が可能である。最近では、話速変換と呼ばれる技術である。音程を変化することなく、話速変換が可能である。

【0007】

すなわち、音声認識において、不特定話者のうち、発声速度に個人差があり、特に早口でしゃべる話者の音声認識率の低下に注目し、話速変換技術を利用して、早口の話者でも音声認識率を向上させようとするものである。

【0008】

【発明が解決しようとする課題】

しかしながら前記従来の構成では、早口でしゃべる話者の音声認識率は向上できるが、男性の低い声、女性、子供の高い声など、幅広い周波数レンジに対応できる、高い音声認識率は実現出来ない。また、早口の場合は、ゆっくり話してもらうなど注意をうながせば済む問題であるが、話者に音色を変えて発声することを望むのは困難である。

【0009】

これは、話者ののどの形状、大きさから、基準の発声周波数が決定されるので、この形状を変えることはできない。また、この音色差による音声認識率の向上を図るには、男性の声、女性、子供の声など音声認識に必要な標準音声データを複数持ち、話者の音色に応じて、参照する標準音声データを切換えなければならない課題を有していた。

【0010】

【課題を解決するための手段】

前記課題を解決するために本発明の音声認識装置は、入力された音声信号の周波数成分を分析し、入力音声を実声認識用の標準音声データに音程変換することで、話者の音色差による音声認識率の向上を図り、しかも標準音声データを複数持つ必要がなく、メモリーの削減を図ることができる。

【0011】

【発明の実施の形態】

本発明の音声認識装置は、入力された音声信号の周波数成分を分析し、入力音声を音声認識データの標準音程に音程変換することで、音声認識率の向上を実現

しうるものである。

【0012】

つぎに、本発明の音声認識装置は、入力された音声信号をアナログ信号からデジタル信号に変換するA/Dコンバータと、前記A/Dコンバータでデジタル信号に変換された入力音声信号を音程変換する音程変換処理部と、前記音程変換処理部で音程変換された音声デジタル信号を基に、音声認識分析する音声認識処理部を直列に具備し、前記音声認識処理部と並列に、音声周波数成分パターンを参照する標準音声データ部を具備すると共に、前記A/Dコンバータの出力を入力とし、入力音声信号の周波数成分を分析する周波数成分分析部と、前記周波数成分分析部の分析結果を入力とし、入力された音声信号の音程をどの程度上下に音程変換するかを判断する音程変換制御部を直列に具備し、前記音程変換制御部で決定された音程変換率で前記音程変換処理部により音程変換を行なう構成とし、音声認識率の向上を実現しうるものである。

【0013】

以下本発明の実施の形態について、図1から図3を用いて説明する。

【0014】

図1は、本発明の一実施例における音声認識装置のブロック構成図を示す。図1において、符号1は入力された音声信号をアナログ信号からデジタル信号に変換するA/Dコンバータ、2は音声信号を音程変換する音程変換処理部、3は音程変換された音声デジタル信号を基に、音声認識分析する音声認識処理部、4は音声認識の音声周波数成分パターンを参照する標準音声データ部、5は入力音声信号の周波数成分を分析する周波数成分分析部、6は周波数成分分析部の分析結果を基に、入力された音声信号の音程をどの程度上下に音程変換するかを判断し、音程変換率を決定する音程変換制御部である。

【0015】

以上のように構成された音声認識装置について、その動作を説明する。図1において、入力された音声信号は、A/Dコンバータ1により、アナログ信号からデジタル信号に変換される。デジタル信号に変換された入力音声信号は、周波数成分を分析する周波数成分分析部5で、高速フーリエ変換により周波数スペクト

ル分析され、図2に示す様な周波数特性例が得られる。同じ一音でも、男性の場合は図2の(a)に示す低域側に周波数スペクトル例が現れ、女性、子供の場合は図2(c)に示す高域側に周波数スペクトル例が現れる。

【0016】

図2(b)は音声認識用の標準音声データとしてメモリーされている標準音声データ部4に格納されている一音の音声周波数スペクトル例を示す。この周波数成分分析部5の分析結果を基に、音程変換制御部6により、音程変換率の算出を行なう。音程変換率は次の様に求める。

【0017】

図2の男性周波数スペクトル例(a)の基本周波数を $f_m$ とし、標準音声データ部4に格納されている標準音声データの周波数スペクトル例(b)の基本周波数を $f_s$ 、女性、子供の周波数スペクトル例(c)の基本周波数を $f_c$ とすると、標準音声データより基本周波数が低い男性では、 $f_s / f_c$ の計算により、1以上の音程変換率が求まり、逆に標準音声データより基本周波数が高い女性、子供の場合は $f_s / f_c$ の計算により、1以下の音程変換率が求まる。

【0018】

音程変換率が1以上の時は音程を上げる方向であり、音程変換率が1以下の時は、音程を下げる方向を意味する。この様にして決定された音程変換率を基に、音程変換処理部2により、音程変換を行なう。音程変換処理は、最近のデジタル信号処理技術を使用すると、簡単に処理可能であり、図3に示すごとく行なう。

【0019】

図3は音声波形の時間変化例であり、図3(a)は音程の低い場合の音声波形の時間変化例であり、図3(b)は標準音声データ部4に格納されている音声波形の時間変化例、図3(c)は音程の高い場合の音声波形の時間変化例であり、図3の(a)を(b)に音程変換するには、入力音声波形をメモリーするときのサンプリングクロックより、速い読み出しクロックで読み出せば実現でき、逆に図3の(c)を(b)に音程変換するには、入力音声波形をメモリーするときのサンプリングクロックより、遅い読み出しクロックで読み出せば実現でき、その読み出しクロックは音程変換率で決定されたクロックとなる。



## 【0020】

但し、この様にすると、音程変換は可能であるが、音程を上げる場合は、音声波形の時間軸が短くなるので、この場合は母音波形を追加して速度変換してやり、音程が下がる場合は、音声波形の時間軸が長くなるので、この場合は、母音波形を間引くのが、一般的な方法として知られている。また、音程変換する為の読み出しクロックは音程変換率を基に、マスタークロックからの分周比で容易に作成できる。

## 【0021】

この様にして音程変換された音声デジタル信号を基に、音声認識分析する音声認識処理部3に入力し、音声認識の音声周波数成分パターンを参照する標準音声データ部4を利用して入力音声周波数成分パターンとの一致度を算出し、音声認識分析する。この様に入力された音声を、あらかじめ、格納された標準音声データの音程に音程変換することにより、標準音声データを複数もつ必要がなく、不特定話者の幅広い周波数レンジにも対応できる音声認識率の向上ができる。

## 【0022】

## 【発明の効果】

以上のように、本発明の音声認識装置は、入力された音声信号の周波数成分を分析し、入力音声を音声認識用の標準音声データに音程変換することで、話者の音色差による音声認識率の向上を図り、しかも標準音声データを複数持つ必要がなく、メモリーの削減を図ることが可能となる長所を有する。

## 【図面の簡単な説明】

## 【図1】

本発明の実施の形態における音声認識装置のブロック構成図

## 【図2】

音声の周波数特性例を示す図

(a) 男性の場合の周波数スペクトル例を示す図

(b) 音声認識用の標準音声データ部4の音声周波数スペクトル例を示す図

(c) 女性、子供の場合の周波数スペクトル例を示す図

## 【図3】

音声波形の時間変化例を示す図

(a) 音程の低い場合の音声波形の時間変化例を示す図

(b) 標準音声データ部 4 に格納されている音声波形の時間変化例を示す図

(c) 音程の高い場合の音声波形の時間変化例を示す図

【図 4】

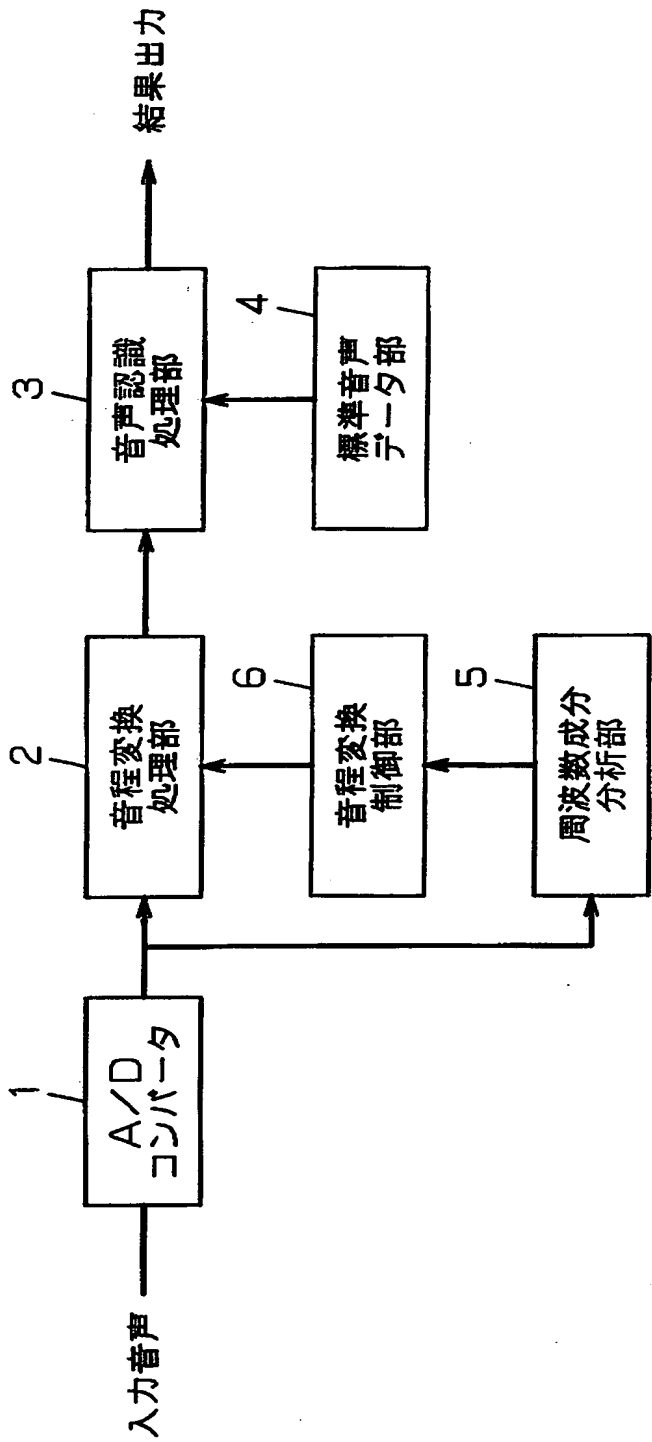
従来の音声認識装置のブロック構成図

【符号の説明】

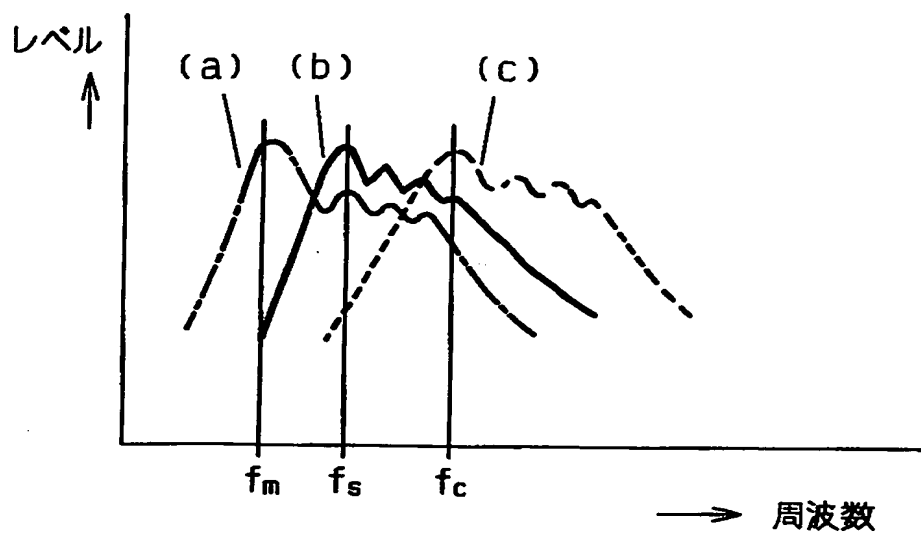
- 1    A/Dコンバータ
- 2    音程変換処理部
- 3    音声認識処理部
- 4    標準音声データ部
- 5    周波数成分分析部
- 6    音程変換制御部
- 1 1   音声入力部
- 1 2   発声速度算出部
- 1 3   発声速度変換率決定部
- 1 4   発声速度変換部
- 1 5   音声認識部

【書類名】 図面

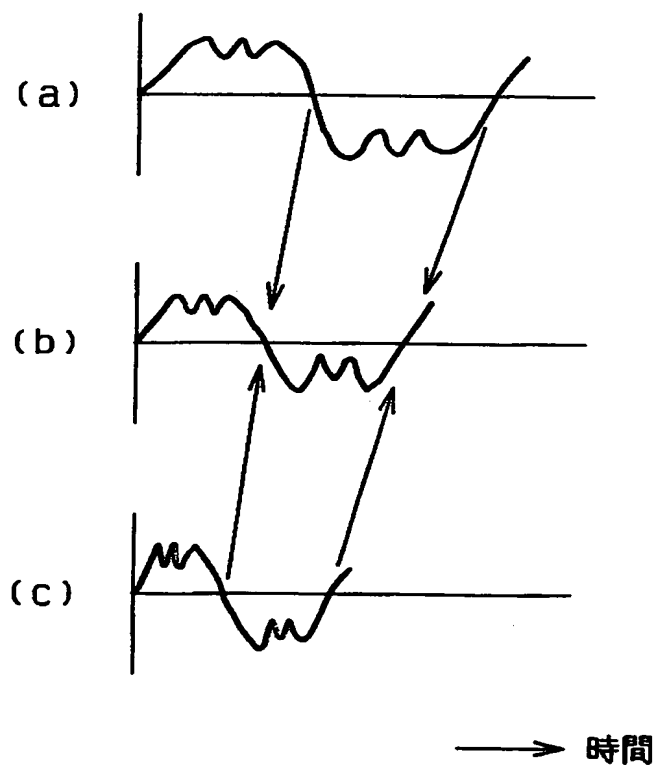
【図 1】



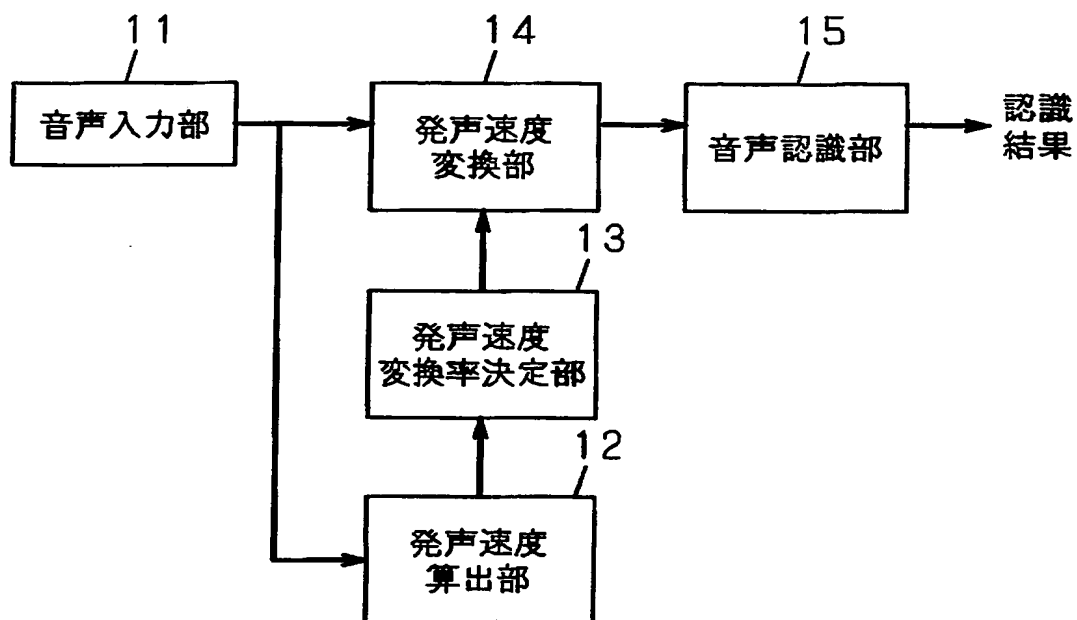
【図2】



【図3】



【図 4】



【書類名】 要約書

【要約】

【課題】 不特定話者の音声認識率の向上を図る。

【解決手段】 入力された音声信号をアナログからデジタルに変換するA/Dコンバータとデジタル信号に変換された入力音声信号を音程変換する音程変換処理部と音程変換された音声デジタル信号を基に、音声認識分析する音声認識処理部を直列に具備し、前記音声認識処理部と並列に、音声周波数成分パターンを参照する標準音声データ部を具備すると共に、前記A/Dコンバータの出力を入力とし、入力音声信号の周波数成分を分析する周波数分析部と、前記周波数成分分析部の結果を入力とし、入力された音声信号の音程をどの程度上下に音程変換するかを判断する音程変換制御部を直列に具備し、前記音程変換制御部で決定された音程変換率で前記音程変換処理部により音程変換を行なう構成とし、不特定話者の音程を標準化することで、音声認識率の向上を図る。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000005821]

1. 変更年月日	1990年 8月28日
[変更理由]	新規登録
住 所	大阪府門真市大字門真1006番地
氏 名	松下電器産業株式会社

*This Page Blank (uspto)*